

MULTIMEDIA



UNIVERSITY

STUDENT ID NO

--	--	--	--	--	--	--	--	--	--

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 1, 2018/2019

BMD3024 – MULTIVARIATE DATA ANALYSIS

(All sections / Groups)

19 OCTOBER 2018
9.00 a.m. - 11.00 a.m.
(2 Hours)

INSTRUCTIONS TO STUDENTS

1. This question paper consists of **SIX (6)** printed pages with **FOUR (4)** structured questions excluding cover page.
2. Marks are shown at the end of each question.
3. Write all your answers in the answer booklet provided.
4. Answer **ALL** questions.

Question One (25 marks)

- (a) Differentiate between the magnitudes of original logistic with the exponentiated coefficients found in the output for logistic regression. (8 marks)
- (b) Compare between Analysis of Variance (ANOVA) and Multivariate Analysis of Variance (MANOVA). (8 marks)
- (c) Describe **TWO (2)** ways that perceptual mapping techniques can be classified into. (9 marks)

Question Two (25 marks)

- (a) Answer the following question based on **Figure 1** below.

Figure 1: Logistic Regression Menu

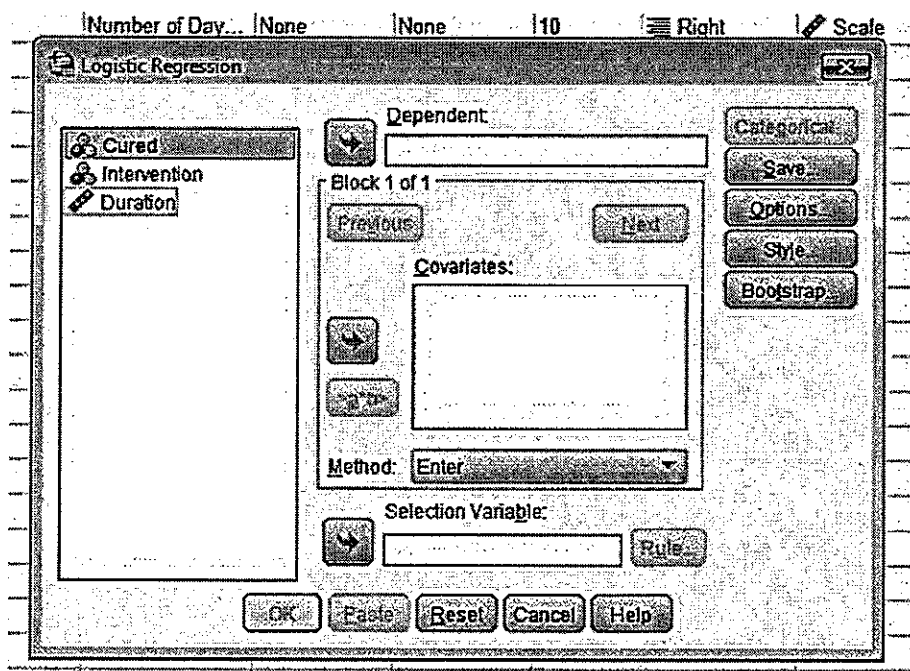


Figure 1 above shows three (3) variables: *Cured*, *Intervention*, and *Duration*. Determine which of the three variables are appropriate for the 'Dependent' and 'Covariates' variables spaces in **Figure 1**.

(8 marks)

- (b) Tables 1 to 3 below shows the survival status of passengers on the HMS Titanic (*survived*, 0 = Died, 1 = Survived), which sank in 1912. The independent variables are age (*age*), and cost of ticket (*fare*). Data was available for 1309 passengers.

Table 1: Classification Table^a

	Observed		Predicted		
			survived		Percentage Correct
			0 Died	1 Survived	
Step 1	survived	0 Died	561	57	90.8
		1 Survived	319	108	25.3
	Overall Percentage				64.0

a. The cut value is .500

Table 2: Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	86.477	2	.000
	Block	86.477	2	.000
	Model	86.477	2	.000

Table 3: Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
Step 1 ^a	age	-.017	.005	12.513	1	.000	.983	.974	.993
	fare	.013	.002	56.059	1	.000	1.014	1.010	1.017
	Constant	-.341	.152	5.063	1	.024	.711		

a. Variable(s) entered on step 1: age, fare.

- (i) Examine Table 1. How accurate is the overall prediction of passengers' survival status on the HMS Titanic?
(2 marks)
- (ii) Examine Table 1. How well do the significant variables predict passengers' that survived the sinking ship?
(2 marks)
- (iii) Examine Table 2. When both independent variables are considered together, do they significantly predict whether a passenger survives the sinking ship? Explain your answer.
(5 marks)
- (iv) Examine Table 3. Which variable(s) make significant contributions ($p < 0.05$) in predicting a passenger survives the sinking ship? Elaborate your answer.
(8 marks)

Continued...

Question Three (25 marks)

- (a) Researchers were interested in the effects of demographic variables, and pre-existing health conditions such as high blood pressure and diabetes on the length and cost of hospital stay. They compared patients based on their gender (*gender*), history of diabetes (*diabetes*), and blood pressure (*bp*). There are two dependent measures: the length of hospital stay (*los*), and treatment costs (*costs*).

Table 4: Box's Test of Equality of Covariance Matrices^a

Box's M	51.097
F	1.450
df1	33
df2	4943.921
Sig.	.046

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

a. Design: Intercept + gender + diabetes + bp + gender * diabetes + gender * bp + diabetes * bp + gender * diabetes * bp

- (i) Refer to **Table 4** above. Using 5 percent level of significance, was the assumption of homogeneity of covariance violated?

(4 marks)

Continued...

Table 5: Multivariate Tests^a

Effect		Value	F	Hypothesis df	Error df	Sig.
Intercept	Pillai's Trace	.724	1929.597 ^b	2.000	1468.000	.000
	Wilks' Lambda	.276	1929.597 ^b	2.000	1468.000	.000
	Hotelling's Trace	2.629	1929.597 ^b	2.000	1468.000	.000
	Roy's Largest Root	2.629	1929.597 ^b	2.000	1468.000	.000
gender	Pillai's Trace	.003	2.060 ^b	2.000	1468.000	.128
	Wilks' Lambda	.997	2.060 ^b	2.000	1468.000	.128
	Hotelling's Trace	.003	2.060 ^b	2.000	1468.000	.128
	Roy's Largest Root	.003	2.060 ^b	2.000	1468.000	.128
diabetes	Pillai's Trace	.002	1.549 ^b	2.000	1468.000	.213
	Wilks' Lambda	.998	1.549 ^b	2.000	1468.000	.213
	Hotelling's Trace	.002	1.549 ^b	2.000	1468.000	.213
	Roy's Largest Root	.002	1.549 ^b	2.000	1468.000	.213
bp	Pillai's Trace	.004	1.366	4.000	2938.000	.243
	Wilks' Lambda	.996	1.366 ^b	4.000	2936.000	.243
	Hotelling's Trace	.004	1.366	4.000	2934.000	.243
	Roy's Largest Root	.003	2.352 ^c	2.000	1469.000	.096
	Hotelling's Trace	.002	.574	4.000	2934.000	.682
	Roy's Largest Root	.002	1.132 ^c	2.000	1469.000	.323

- (ii) Table 5 shows multivariate tests output. Using Hotelling's Trace statistic and 5 percent level of significance, determine whether *gender*, *diabetes*, and *bp* are significant predictors for the dependent variate.

(3 marks)

Continued...

- (b) Cluster analysis was done on a dataset containing **THREE (3)** variables – *work*, *price*, and *salary*, which has 46 respondents. **Table 6** shows the part of the output for Agglomeration Schedule using Ward's Method – Squared Euclidean Linkages.

Copy Table 7 into your Answer Booklet. Complete it based on Table 6.
(18 marks)

Table 6: Agglomeration Schedule

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
36	28	39	4577.340	25	0	37
37	2	28	6669.390	33	36	39
38	4	6	9714.164	35	30	41
39	1	2	12507.427	32	37	43
40	15	25	12581.880	0	0	42
41	4	8	20293.711	38	34	44
42	3	15	31635.600	31	40	45
43	1	32	34843.904	39	0	44
44	1	4	60827.324	43	41	45
45	1	3	173518.080	44	42	0

Table 7: Ward's Method - Squared Euclidean Linkage

No. of clusters	Agglomeration last step	Coefficients this step	Change
2			
3			
4			
5			
6			
7			

Continued...

Question Four (25 marks)

- (a) Develop **ONE (1)** path diagram that shows the following hypothesized structural relationships amongst **THREE (3)** constructs: *X1*, *X2*, and *Y*; and **TWO (2)** mediating variables, *Education* and *Gender*.

- *X1* and *X2* are independent factors.
- *Y* is a dependent factor.
- *X1* is made of five (5) variables.
- *X2* and *Y* are made of four (4) variables.
- *Gender* mediates the relationship between *X1* and *Y*.
- *Education* mediates the relationship between *X2* and *Y*.
- *X2* is also directly related to *Y*.

(20 marks)

- (b) In the path analysis diagram drawn for question (a), label all the following hypotheses and relationships between factors in the diagram:

- H₁: *X1* is positively related to variable *Gender*.
H₂: *Gender* is positively related to *Y*.
H₃: *X2* is positively related to variable *Education*.
H₄: *Education* is positively related to *Y*.
H₅: *X2* is positively related to *Y*

(5 marks)

End of Page